Humans and Machines on Social Media and Fake News: A Review of Two Books

Paul Jason V. Perez
Assistant Professor
University of the Philippines School of Library and Information Studies
pj@slis.upd.edu.ph

RESOURCE REVIEW

Book review of

Bail, C. (2021). *Breaking the social media prism: How to make our platforms less polarizing.* Princeton University Press.

Giansiracusa, N. (2021). *How algorithms create and prevent fake news: Exploring the impacts of social media, deepfakes, GPT-3, and more.* Apress

> *"Fake news is fundamentally an informational problem, and hence it should be a core concern for LIS." — Brendan Luyt (2021)*

During the 10th Asia-Pacific Library and Information Education and Practice (A-LIEP) Conference, Brendan Luyt, one of the invited speakers for the LIS Pedagogy Panel, discussed the link of fake news with human rights, social justice, and equality and why it should be front and center in library and information science (LIS) education and research. Unfortunately, as Luyt has further discussed, the LIS field is not doing enough. As an LIS educator, I have encountered many LIS students who have expressed their interest in studying misinformation and disinformation. I have told them that they are in the right place to do so. However, as Luyt has mentioned, I, too, felt that the LIS field still has a lot to catch up on in studying and researching fake news. Thus, I offer this review of two books, written by professors from sociology and mathematics examining humans and machines on social media platforms to LIS students and researchers of misinformation and disinformation. This review will provide some of the major talking points of the books and argue that the LIS field has a unique opportunity to address the problems discussed in the two resources.

Another reason for people, not just in LIS, to read the two books is that fake news and social media usage are the most significant issues facing society now. Especially for the Philippines, where it is the ground zero for fake news in social media (Ong & Cabañes, 2018).

ABOUT THE AUTHORS

Christopher Bail is a professor of sociology and public policy and founder of the Polarization Lab at Duke University. He is well-known for researching issues on political polarization in the United States, especially on matters of race, religion, and immigration. His most cited article on Google Scholar, written with his colleagues, is a 2018 article that examines what happens when people are exposed to views different from their own (Bail et al., 2018), which has become the basis of his book.

Noah Giansiracusa, on the other hand, teaches mathematics and data science at Bentley University in Waltham, Massachusetts. He has been quoted multiple times in Forbes articles on artificial intelligence (AI) and deepfakes and has taught a course entitled *Seeking Truth with Data. How Algorithms Create and Prevent Fake News* is his first book. Although Giansiracusa comes from a highly technical field of mathematics and data science, his work was praised by many, including a Nobel laureate in economics, for explaining machine learning and AI in simple terms, which is one of the reasons this book was reviewed.

SOCIAL MEDIA AS A PRISM

Perhaps the most notable contribution of *Breaking the Social Media Prism* is the critique of our notion of social media as echo chambers and the argument for a new metaphor which is *social media as a prism*. These

ideas were carefully explained in chapters one to six, which are the most of the book.

According to Bail, the normative assumption about social media users is that they come in and, in time, follow people and ideas similar to and validate their own. Eventually, they become trapped in a feedback loop that reinforces their ideas, known as the echo chamber. This assumption leads to attempts to "break" the echo chamber, such as exposing people to views that are different from their own, hoping that doing so will allow users to reevaluate their beliefs in light of new information. However, as this book has pointed out, the reverse is more likely to happen. Exposing people to ideas that are different from their own pushes them to cling to their beliefs more.

As an alternative, Bail offers that instead of seeing social media as an echo chamber, we can look at it as a prism that distorts reality. In particular, it amplifies the voice of a few users with extreme views on issues while silencing the points of view of the many with moderate and similar opinions. This distortion of reality, or social media prism, is what causes polarization on social media.

These claims were drawn from research conducted by Bail and the Polarization Lab through carefully designed experiments and ethnography of real people using social media. The stories and quotes from their subjects provide a face on an otherwise unhuman platform that we see on social media.

So how do we break the prism? The remaining chapters of the book provided steps on how to do this. Chapter 7 explores the question of what happens when we delete our social media accounts and whether it is genuinely possible to divorce our daily lives from big tech platforms. Chapter 8 starts discussing how we can start seeing ourselves in the prism and break them. Finally, the last chapter offers tools developed by the Polarization Lab that ordinary citizens can use to break the social media prism.

One obvious criticism of Bail's work is that it is American-centric, which is understandable given that one of their missions is to understand American society. The Philippines is unique given that there is evidence pointing to troll farms, social media users deliberately operated by mechanisms to sow misinformation and disinformation (Ong & Cabañes, 2018). Personal communication with the author confirmed that they only focused on human users of the platform and not troll farms.

Nevertheless, the work of Bail and his Polarization Lab offers many insights on how human behaves when confronted with polarizing views and fake news on social media. Moreover, the book provides an appendix of research methods detailing the experiments done through social media, which is very useful for students and researchers.

ECONOMICS OF FAKE NEWS AND ITS RELATION TO AI

In *How Algorithms Create and Prevent Fake News*, Giansiracusa started the book with a chapter on the economics and history of news production and how news content is propagated horizontally and vertically by small and big news organizations. This chapter serves as a foundation in analyzing the incentives of various organizations and corporations to create and share fake news in the following chapters.

Chapters 2 and 3 serve as a crash course on Machine Learning, specifically on methods and algorithms that generate texts, images, and videos. Giansiracusa showed expertise in explaining how the algorithms like Generative Pre-trained Transformer 3 (GPT-3) and Generative Adversarial Network (GAN) work. The author also provided the historical and social contexts of the algorithms. In these chapters, the recurring theme of a race between producing and detecting fake news starts. Some algorithms are tasked to produce texts, images, and videos, while others are created to detect these AI-produced content. One example given in the book is that early deepfake videos feature human beings that blink less. This information was added to algorithms trying to detect deepfake videos, and thus, were able to adjust, making the detection harder.

Chapters 4 to 8 are a closer examination of how the algorithms are deployed in various online platforms. YouTube algorithms for video recommendations were discussed in Chapter 4. Chapter 5 is an interesting chapter with fun trivia on the relationship between Wonder Woman and lie detectors or polygraphs and the not-so-fun reality of attempting to automate lie detection tests. Chapter 6 is perhaps the longest as it discusses the different Google products and the ways in which fake news may propagate through them—from maps, news, images, and search. Chapter 7 continues the discussion with Google but focuses on advertisements with some discussion on Facebook political ads. While Chapter 8 is dedicated to the spread of fake news on social media platforms like Facebook and Twitter, along with the challenges in content moderation. All of the chapters showed that big tech companies had used algorithms to maximize user engagement.

Interestingly, even though Giansiracusa's book came out before the Facebook Whistleblowers (Hao, 2021; Paul & Anguiano, 2021), the points raised, especially in Chapter 8, support the claims that Facebook algorithms amplify hateful content. The bottom line of big tech companies is their profits for their shareholders. As Bail mentioned in *Breaking the Social Media Prism*:

> *Suppose we could identify a small tweak to Facebook's platform that would reduce the number of uncivil exchanges ... by 7.5 percent, though it would also decrease ad clicks by 5 percent. Would the company's leadership and board members go for it? (p. 97)*

The last chapter of Giansiracusa's work provides tools for detecting and minimizing the spread of fake news.

WHY SHOULD LIS STUDENTS AND PROFESSIONALS READ THESE BOOKS?

LIS students and researchers interested in social media, fake news, and AI will benefit from these two books. Not only have they provided a foundational introduction to these concepts, but they have also listed some of the latest studies done in the field. Both resources also showed technical tools that can be used to detect fake news, create network visualization from Twitter data, and identify bot-like accounts.

Perhaps the most important thing that both resources have in common is that both point toward the insufficiency of current AI and machine learning solutions in detecting and preventing fake news. This aspect of social media platforms remains human (Roberts, 2021), especially content moderation (Roberts, 2014). Even if AI and/or machine learning advances in solving fake news, most of this only applies in English-speaking countries, as shown in Chapter 8 of Giansiracusa's work. Misinformation detected in English can still propagate unchecked in the Spanish language; what more for the languages of the developing countries?

Thus, when algorithms fail in identifying truthful and trustworthy information, the role of LIS professionals has never been more needed. Roberts (2014) mentioned that,

> *As the internet has ceded its space to more and more sites of corporatized control and models of information sharing that are fundamentally driven by profit motive before all other values, libraries have remained largely more transparent, more open, and more responsible to the public. (p. 217)*

There is an opportunity for LIS professionals to lead communities in navigating the social media infosphere. This, of course, will require funding, support, and leadership.

DECLARATION ON CONFLICTING INTERESTS

The author declared no potential conflicts of interest with respect to research, authorship, and/or publication of this article.

DECLARATION ON SOURCES OF FUNDING

The author received no financial support for the research, authorship, and/or publication of this article.

REFERENCES

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences, 115*(37), 9216–9221. https://doi.org/10.1073/pnas.1804840115

Hao, K. (2021, August 24). *She exposed how Facebook enabled global political manipulation. Now she's telling her story.* MIT Technology Review. https://www.technologyreview.com/2021/07/29/1030260/facebook-whistleblower-sophie-zhang-global-political-manipulation/

Luyt, B. (2021). LIS Pedagogy Panel. *10th Asia-Pacific Library and Information Education and Practice Conference.* [Video]. YouTube. https://youtu.be/ojrFfcJsmkU

Ong, J. C., & Cabañes, J. V. (2018). *Architects of networked disinformation: Behind the scenes of troll accounts and fake news production in the Philippines.* https://doi.org/10.7275/2cq4-5396

Paul, K., & Anguiano, D. (2021, October 23). *Facebook crisis grows as new whistleblower and leaked documents emerge.* The Guardian. https://www.theguardian.com/technology/2021/oct/22/facebook-whistleblower-hate-speech-illegal-report

Roberts, S. T. (2014). *Behind the screen: The hidden digital labor of commercial content moderation.* University of Illinois at Urbana-Champaign.

Roberts, S. T. (2021). Your AI is a human. In T. S. Mullaney, B. Peters, M. Hicks, & K. Philip (Eds.), *Your computer is on fire.* MIT Press.

AUTHOR BIOGRAPHY

Paul Jason V. Perez is a full-time faculty member of the UP School of Library and Information Studies where he teaches introductory courses on LIS, ICT, and Programming. Before joining the school, he worked as a Senior Technical Associate for UP Information Technology Development Center, where he was involved in implementing an information system across the different campuses of the university.

He finished his Bachelor of Library and Information Science from the same school. In 2019, he completed his Master of Digital Information Management from the University of Technology Sydney through the Australia Awards Scholarship. His current research interests include Digital Libraries, Linked Open Data, Digital Assets Management, Research Data Management, among others. His latest projects, talks, and publications can be accessed on his website at *https://pjperez.xyz*.

The PhJLIS is published by the
School of Library and Information Studies,
University of the Philippines Diliman.
ISSN 2719-0471 (Online)